

10/501857

## Marked-Up Version of Substitute Specification

## Description

~~Language recognizer and operating method therefor~~APPARATUS AND METHODS FOR LANGUAGE RECOGNITION

5

BACKGROUND

The present ~~disclosures invention~~ relates to a voice recognizer ~~according to the preamble of claim 1 and an operating method therefore, that stores a vocabulary of words to be recognized for voice control of a plurality of programs and/or other files, each of which is assigned a word of the vocabulary as a name.~~

10

Having long secured itself a permanent and constantly growing application area in the input of text to office applications running on PCs, voice recognition is also making increasing inroads in the control of technical devices. Both in ultra-

15 miniaturized and at the same time computerized hand-held electronic devices, in particularly mobile phones and PDAs, and in technical devices that are meant to involve minimum attention and concentration from the user to operate, such as the various technical devices in a moving car, this type of voice recognition together with voice control based thereon can find useful potential applications. In the

20 former type of devices, the area available for control actions has actually become so small that the numerous possible functions can only be implemented very inconveniently using traditional keypad or touch-screen entries, (and almost not at all for people with poor sight). In areas of use in which the attention of the user must remain focused on other things, (e-g-for example road traffic), the introduction

25 of voice control not only increases convenience but greatly improves safety.

In voice recognition, a lexicon containing the words to be recognized is required. In the case of phoneme-based voice recognition, these words are transferred by means of a text-to-phoneme technique into a phonetic transcription

30 and saved in the vocabulary. During the recognition process, a search for the best path through the phoneme strings contained in the vocabulary is made using the

Viterbi algorithm as it is known. Details of the established voice recognition algorithms are given in the relevant technical literature.

Highly computerized technical devices of the aforementioned type (for example PDAs, hand-held PCs, mobile phones, vehicle audio systems, on-board computers, etc.) have user-interfaces or MMI structures that are derived from PC user interfaces. There are a large number of applications installed that need to be controlled in a suitable way, and also in more complex devices in a specific sub-level of a logical hierarchy. In traditional devices of this type, menu-based control is provided for this purpose that can be executed by the user using soft-key entries.

When selecting an application by voice input, the program names of the available applications are contained in the lexicon. Once a name is recognized, the relevant program is executed or the application started. To do this, the program name and the program path must be saved in a suitable format.

According to the state of the art, the individual program names are hard-wired to the corresponding recognition results (the words in the lexicon). This can be specified in an additional file, or permanently defined in the source code of the program. Both methods have ~~essential~~ disadvantages, which are described below:

—One disadvantage is that when working with an additional file there is the problem that it can be seen by the user and consequently can also be modified. Even binary formats or write-protected files offer no effective protection against changes. This can lead to discrepancies between the vocabulary used and the word list or program list, with the consequence that the application may respond incorrectly.

—Another disadvantage is that when the voice expressions acting as control commands are defined in the source code, it is not easy to make further

changes to the vocabulary. The source code would need to be re-compiled and shipped every time changes in the program names occurred.

~~The crucial~~ Additionally, a further disadvantage of the technique used up to now is the non-existent or inadequate system expandability. At present, it is not possible for the user to record his own commands or applications for inclusion in the automatic voice recognition, at least not without the risk of a fault in the originally programmed configuration of the voice recognizer.

## 10 SUMMARY OF THE INVENTION

The ~~object of the present disclosure invention is thus to~~ provides an improved voice recognizer and methods for its operation with which ~~said the~~ device can be configured more flexibly in order to include the user's own control commands or applications.

15

As an example, an apparatus for voice recognition is provided including This object is achieved in its device aspect by a voice recognizer having the features of claim 1, and in its a storage having a stored vocabulary of words to be recognized for voice control of a plurality of programs and other files, wherein each of the plurality of programs and other files is assigned a word of the vocabulary as a name. The apparatus also includes a file directory configured to store a link to each program and file of the plurality of programs and other files, wherein the names of the links form a first active partial vocabulary of the voice recognition apparatus.

20

As another example, a voice recognition method is provided comprising aspect by an operating method having the features of claim 6, providing a voice recognition apparatus, for example, as described above, and generating a current vocabulary containing at least the names of the links from the file directory when a voice recognizer program configured to perform voice recognition is started.

25

## DETAILED DESCRIPTION OF THE PRESENT EXAMPLES

The ~~invention includes~~ presently disclosed apparatus and methods incorporate the fundamental idea of providing a user interface constructed using links for the voice control of applications or for suitable handling of files. The organization principle of the links enables programs or files in different hierarchy levels to be opened easily in a structured way without a rigid assignment needing to be defined and programmed in advance.

The list of words to be recognized (the lexicon) is defined by the contents of a specific file directory which contains links (shortcuts) to the programs or files present. The name of the link specifies the word to be recognized, and the program or file to which this link points specifies the action to be performed. In converting the name, one should note that only the partial string in front of the first dot is used as a command. The vocabulary is generated when the recognizer program is started. This allows a flexible response to changes in the application structure or file structure. As soon as a word is recognized, the relevant link is actuated and the required action executed.

Advantages compared with ~~the previous techniques used up to now lie in~~ the include flexibility regarding words and actions, and the simple creation and modification of a complex recognizer vocabulary. New commands can be added to the existing vocabulary in a simple and familiar way. A shortcut to the required program or file merely needs to be created in the file directory. Under Windows, for example, a shortcut can be created easily via the context menu.

~~This illustrates a~~ A further advantage ~~:- since of the presently disclosed apparatus and methods is that~~ the file system takes over the management of commands and actions (name and destination of the shortcut), and, therefore, no additional program is required for managing the vocabulary. If a command is meant to be deleted, the link is simply deleted.

Since modern operating systems allow links to files as well, documents can also be opened by voice command.

5 In ~~a preferred embodiment~~ an example, the file directory includes a plurality of sub-directories in at least one subordinate hierarchy level, the directory names forming a first and, if applicable, further, active partial vocabularies of the voice recognizer lower down the hierarchy.

10 By using sub-directories in the file directory, structured voice commands to open programs and files can be generated in the simplest way. For instance, all links to pieces of music are saved in a sub-directory "music". The word "music" is held in the active vocabulary in the first stage of recognition. If it is recognized, the vocabulary is switched (e.g., by language model), and the links contained in the "music" sub-directory are now held in the active vocabulary.

15 In particular, each program or file is assigned from a sub-directory a voice command ~~composed of~~ having multiple connected parts that contains the names of the links from the file directory and each subordinate sub-directory leading to the program or file.

20 Complex voice commands can be created and edited in the simplest way using this method. Existing directories containing shortcuts, such as the Windows start menu, can now be operated simply by voice control because all necessary information is already there.

25 This method is a further development of shortcuts to programs (for example, Windows PC) and the hard-wired voice recognizer resources. In this method the recognizer resource is provided automatically by the creation of a link, i.e. the name of the link can be processed by the recognizer immediately  
30 afterwards.

In general, any files and programs can be opened by voice command once they have been copied into the special directory. It also makes no difference whether a music title, c++ file, text document or program is involved. By saving a link in the special directory, the file is opened by the default program configured.

5 For example, a document with the .doc extension is opened automatically by the Word program (as when double clicking on the file in traditional PC entry).

The aspects of the ~~invention explained above~~above disclosure appear as both ~~as the device apparatus~~aspects of a voice recognizer and as aspects of ~~the an~~an operating method thereof, particularly since it is typically implemented in a suitable mix of hardware and software components.

10

Two ways of recording a word in the recognizer lexicon are ~~given described below~~as follows. ~~(1)The first way is R~~(1)The first way is Recording by a program call via the context menu for the required application. In this case the context menu contains two program calls (e.g., Add and Remove). Add adds the relevant program/file and Remove displays the list of programs/files that can currently be selected by voice selection. ~~(2)The second way is to use~~Using a “drag’n’drop” procedure to copy the link to the required application into the special folder. In this case, in order to

15

20 remove a program, one must switch to the relevant directory and delete the required link from the directory by “deleting”.

~~The implementation of the invention is not limited to the examples and aspects described above, but is possible in numerous variations falling within the bounds of proper action.~~ It should be understood that various changes and modifications to the presently preferred examples described herein will be apparent to those skilled in the art. Such changes and modifications can be made without departing from the spirit and scope of the present invention and without diminishing its intended advantages. It is therefore intended that such changes and

25

30 modifications be covered by the appended claims.

## ABSTRACT OF THE DISCLOSURE

Disclosed is a language recognition apparatus ~~with having a storage with~~  
a stored vocabulary of words to be recognized for the language-based controlling of  
a ~~plurality of~~ programs and/or other files. A word of ~~said the~~ vocabulary is  
5 assigned to each program and/or file. A link is stored in a file directory for each  
program or file, and the names of ~~said the~~ links forming a first active partial  
vocabulary of the language recognition apparatus. Also disclosed is methods of  
operating the language recognition apparatus including providing a language  
recognition apparatus, for example, as described above, and generating a current  
10 vocabulary containing at least the names of the links from the file directory when a  
voice recognition program configured to perform voice recognition is started.

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of claims:**

Claims 1-9 (canceled).

5        Claim 10 (new): A voice recognition apparatus comprising:  
a storage having a stored vocabulary of words to be recognized for voice control of  
a plurality of programs and other files, wherein each of the plurality of programs  
and other files is assigned a word of the vocabulary as a name; and  
a file directory configured to store a link to each program and file of the  
10       plurality of programs and other files, wherein the names of the links form a first  
active partial vocabulary of the voice recognition apparatus.

Claim 11 (new): The voice recognition apparatus as defined in claim 10,  
wherein the names of the links are formed by voice commands, and the links define  
15       shortcuts to application programs.

Claim 12 (new): The voice recognition apparatus as defined in claim 10,  
wherein the names of the links are formed by voice commands, and wherein the  
links define shortcuts to files selected from a group consisting of text documents,  
20       voice documents, music files, and video files.

Claim 13 (new): The voice recognition apparatus as defined in claim 10,  
wherein the file directory contains a plurality of sub-directories in at least one  
subordinate hierarchy level, wherein names of the plurality of sub-directories  
25       together with the names of the links form a first active partial vocabulary of the  
voice recognition apparatus lower down the hierarchy.

Claim 14 (new): The voice recognition apparatus as defined in claim 13,  
wherein each program and file of the plurality of programs and other files is  
30       assigned from at least one of the plurality of sub-directories a voice command  
comprising multiple connected parts that contain the names of the links from the



file directory and the at least one of the plurality of sub-directories leading to the program or file.

Claim 15 (new): A voice recognition method comprising:

5 providing a voice recognition apparatus including a storage having a stored vocabulary of words to be recognized for voice control of a plurality of programs and other files, wherein each of the plurality of programs and other files is assigned a word of the vocabulary as a name, and a file directory configured to store a link to each program and file of the plurality of programs and other files, wherein the  
10 names of the links form a first active partial vocabulary of the voice recognition apparatus; and

generating a current vocabulary containing at least the names of the links from the file directory when a voice recognizer program configured to perform voice recognition is started.

15

Claim 16 (new): The voice recognition method as defined in claim 15, further comprising:

effecting administration of the vocabulary by managing the file directory and at least one sub-directory without an additional vocabulary management  
20 program.

Claim 17 (new): The voice recognition method as defined in claim 16, further comprising:

creating sub-directories below the file directory in at least one subordinate  
25 hierarchy level in order to edit voice commands having multiple connected parts; and

recognizing the voice commands having multiple connected parts in a multi-stage recognition process, wherein in the course of recognizing a switch is made from a first active partial vocabulary into an at least second active partial  
30 vocabulary.

Claim 18 (new): The voice recognition method as defined in claim 17,  
further comprising:

5        recording new words in the vocabulary by effecting a program call via a  
context menu for a relevant program or file of the plurality of programs and other  
files.

Claim 19 (new): The voice recognition method as defined in claim 17,  
further comprising:

10        removing words from the vocabulary by effecting a program call via a  
context menu for a relevant program or file of the plurality of programs and other  
files.

Claim 20 (new): The voice recognition method as defined in claim 17,  
further comprising:

15        recording new words in the vocabulary by effecting a “drag’n’drop”  
procedure.

Claim 21 (new): The voice recognition method as defined in claim 17,  
further comprising:

20        removing words from the vocabulary by effecting a “drag’n’drop”  
procedure.

## REMARKS

The present amendment makes editorial changes and corrects typographical errors in the specification, which includes the Abstract, in order to conform the specification to the requirements of United States patent practice. No new matter is added. Included is a Substitute Specification including a marked-up version of the changes made via by the present amendment.

In addition, the present amendment cancels original claims 1-9 in favor of new claims 10-21. Claims 10-21 have been presented because revisions by red-lining and underlining which would have been necessary in claims 1-9 in order to present this claim in accordance with preferred United States patent practice would have been too extensive, and, thus, too burdensome. The present amendment is intended for clarification purposes only and not for substantial reasons related to patentability pursuant to 35 U.S.C. §§101, 102, 103 or 112. Indeed, the cancellation of claims 1-9 does not constitute an intent on the part of the Applicants to surrender any of the subject matter of claims 1-9.

Early consideration on the merits is respectfully requested.

Respectfully submitted,

BELL, BOYD & LLOYD LLC

BY

  
Patrick B. Law

Reg. No. 41,549

P.O. Box 1135

Chicago, Illinois 60690-1135

Phone: (312) 781-6801

Dated: July 16, 2004